

Testing Relationships between Multiple Regional Features and Biogeographic Processes of Speciation, Extinction, and Dispersal

SARAH K. SWISTON^{*},  AND MICHAEL J. LANDIS^{*}, 

Division of Biology and Biomedical Sciences, Washington University in St. Louis, Rebstock Hall, St. Louis, MO 63130, USA

*Correspondence to be sent to: Department of Biology, Washington University in St. Louis, Rebstock Hall, St. Louis, MO 63130, USA;

E-mail: [\(S.K.S.\); michael.landis@wustl.edu](mailto:sarah.k.swiston@wustl.edu) (M.J.L.)

Received 20 June 2023; reviews returned 17 October 2024; accepted 03 November 2024

Associate Editor: Richard Ree

Abstract.—The spatial and environmental features of regions where clades are evolving are expected to impact biogeographic processes such as speciation, extinction, and dispersal. Any number of regional features (such as elevation, distance, area, etc.) may be directly or indirectly related to these processes. For example, it may be that distances or differences in elevation or both may limit dispersal rates. However, it is difficult to disentangle which features are most strongly related to rates of different processes. Here, we present an extensible Multifeature Feature-Informed GeoSSE (MultiFIG) model that allows for the simultaneous investigation of any number of regional features. MultiFIG provides a conceptual framework for incorporating large numbers of features of different types, including categorical, quantitative, within-region, and between-region features, along with a mathematical framework for translating those features into biogeographic rates for statistical hypothesis testing. Using traditional Bayesian parameter estimation and reversible-jump Markov chain Monte Carlo, MultiFIG allows for the exploration of models with different numbers and combinations of feature-effect parameters and generates estimates for the strengths of relationships between each regional feature and core process. We validate this model with a simulation study covering a range of scenarios with different numbers of regions, tree sizes, and feature values. We also demonstrate the application of MultiFIG with an empirical case study of the South American lizard genus *Liolaemus*, investigating 16 regional features related to area, distance, and elevation. Our results show two important feature-process relationships: a negative distance/dispersal relationship and a negative area/extinction relationship. Interestingly, although speciation rates were found to be higher in Andean versus non-Andean regions, the model did not assign significance to Andean- or elevation-related parameters. These results highlight the need to consider multiple regional features in biogeographic hypothesis testing. [Historical biogeography; *Liolaemus*; phylogenetics; SSE model.]

Biogeographic processes such as speciation, extinction, and dispersal are expected to be shaped by the spatial context in which these processes operate. Any number of regional features—such as those relating to region size, distance, elevation, climate, barrier-presence, or status as an island or continent—may directly or indirectly shape the biogeography of a particular clade. For example, an inverse relationship between island size and probability of extinction is an essential feature of MacArthur and Wilson's theory of island biogeography (MacArthur and Wilson 1967), and region area has since been frequently implicated to explain the patterns of diversification (Connor and McCoy 1979), particularly when historical areas are considered (Fine and Ree 2006). Similarly, an inverse relationship between distance and dispersal has been established, both in terms of physical distance (MacArthur and Wilson 1967) and traversability of connecting habitat (Adriaensen et al. 2003). It has also been suggested that climatic features influence diversification dynamics in a wide variety of clades, including plants (Scheiner and Rey-Benayas 1994), mammals (Buckley et al. 2010), amphibians and reptiles (Navas 2002), and many others.

Many different phylogenetic models of historical biogeography have been created for the purposes of explaining how species both evolve over time and move

through space (Ronquist 1997; Ree et al. 2005; Sanmartín et al. 2008; Goldberg et al. 2011; Matzke 2014; Valente et al. 2015; Albert et al. 2017; Skeels and Cardillo 2018; Quintero and Landis 2020; Landis et al. 2021, 2022). These models vary widely in function and application. However, even with the quantity and breadth of available models, testing hypotheses about the relationships between regional features and evolutionary/biogeographic rates remains a challenge. Due to its inference capabilities and incorporation of state-dependent diversification, the GeoSSE model (Goldberg et al. 2011) provides a good foundation for testing regional hypotheses of historical biogeography (Ree and Sanmartín 2009, 2018; Crisp et al. 2011). The GeoSSE model is based on previous state-dependent speciation and extinction (SSE) models, primarily the Binary State Speciation and Extinction (BiSSE) model (Maddison et al. 2007). Previous SSE models such as BiSSE were used for character evolution rather than biogeography, but the two subjects have many similarities (species range could be considered an evolving character). The GeoSSE framework, which has generally only been applied to biogeographic systems with two regions, assigns one free parameter to every possible speciation, extinction, and dispersal rate and, because of this, we may be able to link the features of those regions to

evolutionary and biogeographic rates. Then, one can jointly examine diversification/biogeographic processes and the individual effects of features on those processes. Unfortunately, the number of free parameters in a standard GeoSSE analysis explodes as the number of regions increases, complicating how we test relationships between regional features and biogeographic event rates. The Feature-Informed GeoSSE (FIG) model (Landis et al. 2022) addresses this issue by modeling biogeographic event rates as parametric functions of regional features. This parameterization strategy both collapses model complexity into something manageable while also creating a framework to test hypotheses about the importance of certain features on evolutionary processes. For example, the original FIG study found that distance and water barriers decreased dispersal rates and increased between-region speciation rates in *Anolis* lizards, as one might expect.

Because of the number of possible regional features (abiotic and biotic) that might influence biogeographic rates, it can be difficult to determine which features impact which processes. This is a major obstacle for hypothesis testing, since the quality of our phylogenetic inferences about the past (biogeographic rates and ancestral states) rely on our ability to choose the best-fitting models. Statistical model selection in historical biogeography has been useful for a wide variety of problems (Ree and Smith 2008; Sanmartín et al. 2008; Landis et al. 2013; Matzke 2014; Cardillo et al. 2017; Landis et al. 2022; Quintero et al. 2023), but typically, the procedure only compares models from a small space of low-complexity models. For example, it could compare a model that includes a distance-dependent dispersal rate parameter with a model that does not. However, we expect that a large number of regional features may influence dispersal rates, such as distances, water-barriers, elevation differences, climatic differences, etc. Testing for the effects of all combinations of features on evolutionary rates is challenging because of the vast model space that must be explored—roughly 2^N model comparisons for N possible feature-process relationship parameters that can be “on” or “off.” One approach to this problem has been the use of generalized linear models with indicators to control model complexity (Lemey et al. 2012; Faria et al. 2013). This method estimates the strengths of feature-process relationships while simultaneously performing model selection via Bayesian stochastic search variable selection (BSSVS). It also addresses the complication of correlated environmental predictors. However, these features only inform rates of dispersal; the model does not incorporate state-dependent diversification and extinction.

Bayesian phylogenetic methods have often overcome the problem of finding sets of good-fitting models using reversible-jump Markov chain Monte Carlo (RJMCMC) (Green 2003; Huelsenbeck et al. 2004). One of the simplest RJMCMC algorithms allows the Markov chain to move among a family of nested models by turning

certain parameters “on” or “off” while also exploring parameter space like a traditional MCMC (akin to the BSSVS method mentioned above). Landis et al. (2022) used RJMCMC to fit FIG models to simulated data and an empirical data set for *Anolis* to control model complexity and to test the combinations of biogeographic hypotheses involving just four regional features: region size, distance, water-barrier presence, and insular region status.

In this paper, we introduce MultiFIG, a scalable Bayesian phylogenetic framework to efficiently explore a large set of biogeographic models for which any number of regional features may or may not change different biogeographic rates of speciation, extinction, and/or dispersal. MultiFIG builds upon the previous FIG model, which used a similar method to investigate hypotheses about the diversification of *Anolis* lizards. The original FIG model was applied to *Anolis* lizards and hand-coded system-specific interdependent relationships between quantitative and categorical feature effects for a small set of regional features (i.e. categorical effects amplified quantitative effects Landis et al. 2022). The MultiFIG framework extends this model by allowing for the inclusion of any number of quantitative and categorical features and relaxes relationships among features. MultiFIG also incorporates essential software improvements and provides a set of tools for interpreting the results of MCMC output when simultaneously testing large numbers of hypotheses. Under the MultiFIG framework, one can incorporate any set of regional features and assess the importance of those features for informing rates of speciation, extinction, and dispersal under models of varying complexity. In this paper, we will present models in which features act independently to inform processes, and demonstrate MultiFIG’s ability to handle a large number of possible feature effects. However, the framework can also allow users to define more complex relationships between features or feature effects. This scalability means that MultiFIG can be used to both 1) identify potential features of importance and 2) test specific, detailed hypotheses.

Using simulated data sets, we apply MultiFIG to a variety of historical biogeographic scenarios to characterize the circumstances for which our hypothesis testing framework performs well or poorly. These scenarios vary in the number of regions, number of taxa, and values/variances of geographic features. We also explore various forms of model misspecification related to the regional features informing the analysis. By examining this breadth of simulation scenarios, we are able to assess model performance under a variety of conditions that may resemble different biological systems where taxa are evolving.

Finally, as a proof-of-concept, we apply the MultiFIG model to an empirical study system: *Liolaemus*, a diverse genus of 199 lizard species that occur primarily within the central and southern Andean highland

and the flanking lowland regions. Current estimates place the origin of *Liolaemus* between 20 and 30 Ma (([Esquerre et al. 2019, 2021](#))), making its diversification concurrent with older episodes of Andean mountain-building ([Gregory-Wodzicki 2000; Garzione et al. 2008; Boschman 2021](#)). This raises the question of whether and how the uplift of the Andes has influenced the evolutionary dynamics and spatial distribution of *Liolaemus*. Interestingly, the genus contains both viviparous and oviparous species, and reconstructions show multiple origins of viviparity across the clade, along with possible reversions to oviparity. Parity mode has a strong correlation with temperature; live-birth is favored in colder climates, and viviparous species have lower active body temperatures ([Esquerre et al. 2019; Ibargüen-goytía et al. 2021](#)). This association suggests that montane habitats may be sites of ecological opportunity for *Liolaemus* lizards, promoting their diversification ([Alencar et al. 2024; Domínguez-Guerrero et al. 2024](#)). Other research suggests that the Andes may contribute to diversification through different mechanisms, such as complex topology facilitating speciation in a changing environment ([Skeels et al. 2023](#)).

Previous work on *Liolaemus* by [Esquerre et al.](#) used a GeoSSE model to investigate how the Andes contributed to evolutionary rates within the clade, finding that speciation rates were higher in Andean than non-Andean regions ([Esquerre et al. 2019](#)). However, the GeoSSE model only allowed for the inclusion of two regions (Andean vs. non-Andean) and was not designed to untangle the effects of many regional features on diversification and dispersal. Therefore, we apply MultiFIG to *Liolaemus*, demonstrating how our framework can uncover more refined relationships between regional features and biogeographic rates.

MATERIALS AND METHODS

Here, we describe an extensible multifeature Feature-Informed GeoSSE (MultiFIG) model. We perform a simulation experiment to validate the model and explore its behavior under different conditions, focusing on different values and variances of geographic features, numbers of regions, numbers of taxa, and types of model misspecification. Then, we apply the model to investigate how regional features related to area, distance, and elevation have impacted the evolution and historical biogeography of *Liolaemus*.

Model Definition

Discrete biogeographic ranges.—The Feature-Informed GeoSSE (FIG) model ([Landis et al. 2022](#)) is an extension of the GeoSSE model ([Goldberg et al. 2011](#)), which is a member of a class of SSE models ([Maddison et al. 2007](#)). SSE models assign lineages to certain “states” that correspond to some attribute they possess, such as a morphological character or a region of residence. Based

on the state of a lineage, it experiences different rates of speciation and extinction, as well as rates of transition to other states. The GeoSSE model specifically handles cases where the states represent possible geographic ranges, comprised of one or more discrete regions inhabited by a species. Take, for example, a clade evolving among $N = 3$ discrete regions (A , B , and C). The subset of regions where a lineage is present defines its range, and its “state” within the SSE framework. In the case of a 3-region system, a species can occupy any of $2^N - 1 = 7$ possible geographic states at a given time: : $\{A\}$, $\{B\}$, $\{C\}$, $\{A, B\}$, $\{A, C\}$, $\{B, C\}$, or $\{A, B, C\}$. Species transition between these states through the processes of dispersal, local extinction (extirpation), and speciation (cladogenesis). Cladogenetic events may directly result in range changes (i.e., daughter lineages do not necessarily possess the same state as the parent). Therefore, the model incorporates four core processes of diversification through space and time: within-region speciation, between-region speciation, extirpation, and dispersal ([Fig. 1](#) depicts GeoSSE dynamics in a 2-region system).

Model events and rates.—The model is constrained such that a lineage can only experience a single event at any instant in time. Dispersal and extirpation are anagenetic processes, operating between cladogenesis events along the branches of an evolutionary tree. For example, dispersal from A to C may cause an original range to expand from $\{A, B\}$ to $\{A, B, C\}$. This allows us to define rates of dispersal events $r_d(i, j)$ for each pair of regions, which can be summed to produce rates of dispersal from any currently inhabited region into a new region. The extirpation rate $r_e(i)$ represents the rate at which a species loses region i from its range. A lineage experiences global extinction when it is extirpated from the final occupied region in its range. For example, extirpation in region A may cause an original range to shrink from $\{A, B\}$ to $\{B\}$. If that species was then extirpated in region B , the lineage would be globally extinct and no longer diversify or disperse.

Within- and between-region speciation are cladogenetic processes, describing different range-inheritance scenarios at the nodes of an evolutionary tree. Let R be the ancestral range, and let S and T be the ranges of the two daughter lineages. Within-region speciation occurs within a single region in a species range, producing one daughter that inherits the parent’s entire range and one daughter with a range consisting of a single region. In a within-region speciation event, $R = S$ and $T \in S$. For example, an ancestral range $R = \{A, B\}$ may yield species with ranges $S = \{A, B\}$ and $T = \{A\}$. Note that both S and T will always be equal to R when the ancestral range consists of only one region. We define rates of within-region speciation events $r_w(i)$ for each region i . For lineages with widespread ranges comprised of multiple regions, a speciation event may also occur between regions, producing two daughters that inherit nonoverlapping subsets of the parent’s range.

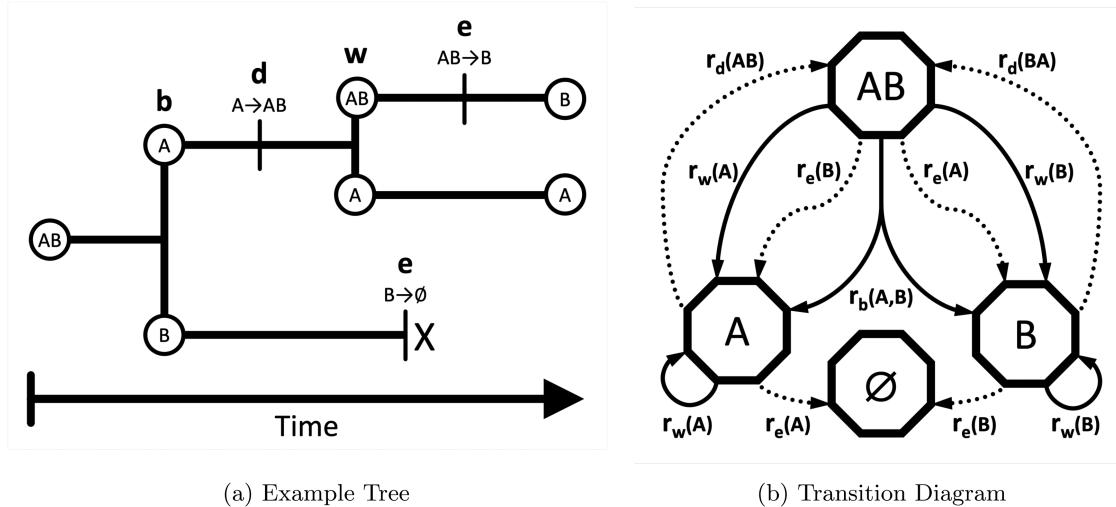


FIGURE 1. (a) Example phylogeny showing all four GeoSSE event types involving a biogeographic system with two regions (A and B): within-region speciation [w], between-region speciation [b], extinction [e], and dispersal [d]. (b) Transition diagram showing event types associated with transitions between different states in a two-region system (GeoSSE model), following the structure of Goldberg et al. (2011) but using FIG notation. Cladogenetic events are shown using solid arrows, while anagenetic events are shown with dashed arrows.

This event is only valid for ancestral ranges consisting of two or more regions. In a between-region speciation event, $R = S \cup T$ and $S \cap T = \emptyset$. For example, an ancestral range $R = \{A, B, C\}$ may split into daughters $S = \{A, B\}$ and $T = \{C\}$. This allows us to define rates of between-region speciation events $r_b(S, T)$, which are related to sets of region pairs (i, j) through a range split score to prevent explosive model overparameterization (Landis et al. 2022).

The parameters associated with these four processes are functions of individual regions, so the evolutionary rates experienced by a lineage are functions of the regions where that lineage is present (and their features). This results in four rate parameters for each region i or region pair i, j : $r_e(i)$ for extinction, $r_w(i)$ for within-region speciation, $r_d(i, j)$ for dispersal, and $r_b(S, T)$ for between-region speciation.

MultiFIG rates.—In the basic GeoSSE model, all regions or region pairs are assigned their own independent parameters representing the core processes, which are dissociated from any regional features that may be causing or correlated with rates variation. This is tractable for a scenario with few regions (the creators of the GeoSSE model applied a system with $N = 2$ regions), but it presents several problems when the number of regions is large. Exploring a large state space (all possible region combinations) makes it challenging to calculate likelihoods, causing performance degradation. We use TensorPhylo to address this issue (May and Meyer 2022). Also, estimating so many distinct rates means estimating large numbers of distinct event rate parameters (see Landis et al. (2022) Supplementary Table S2). Because of this, we may not have sufficient data to estimate these

parameters with any accuracy. Another problem is that large numbers of free parameters create a highly multimodal posterior surface, which means that Bayesian inferences such as MCMC may mix slowly, leading to high computational costs. This leaves many questions with large state spaces and/or complex biogeographic model structures understudied. However, if we expect that certain features shape biogeographic rates in similar ways across regions, we can express evolutionary rates as functions of the features of the regions where lineages are evolving. For instance, we may hypothesize that dispersal between two regions scales with distance, and we can therefore represent dispersal as a function of the data (known distances between regions) and a single parameter describing the strength and direction of the effect of distance on dispersal. This allows us to reduce the number of model parameters and to directly examine whether particular region features correlate with increased or decreased rates of certain processes. The Feature-Informed GeoSSE (or FIG) model (Landis et al. 2022) presents a way to parameterize the combined effects of many region features, allowing for inference of biogeographic history, evolutionary rates, and model parameters in scenarios with numerous regions.

The FIG model incorporates geographical features with two value types as model variables: quantitative features and categorical features. Quantitative features have continuous real values, while categorical features have discrete values. In our analyses, we only used positive-valued quantitative features and binary-valued categorical features, though these assumptions can be relaxed to suit the needs of the research question. The two value types are handled differently by the model,

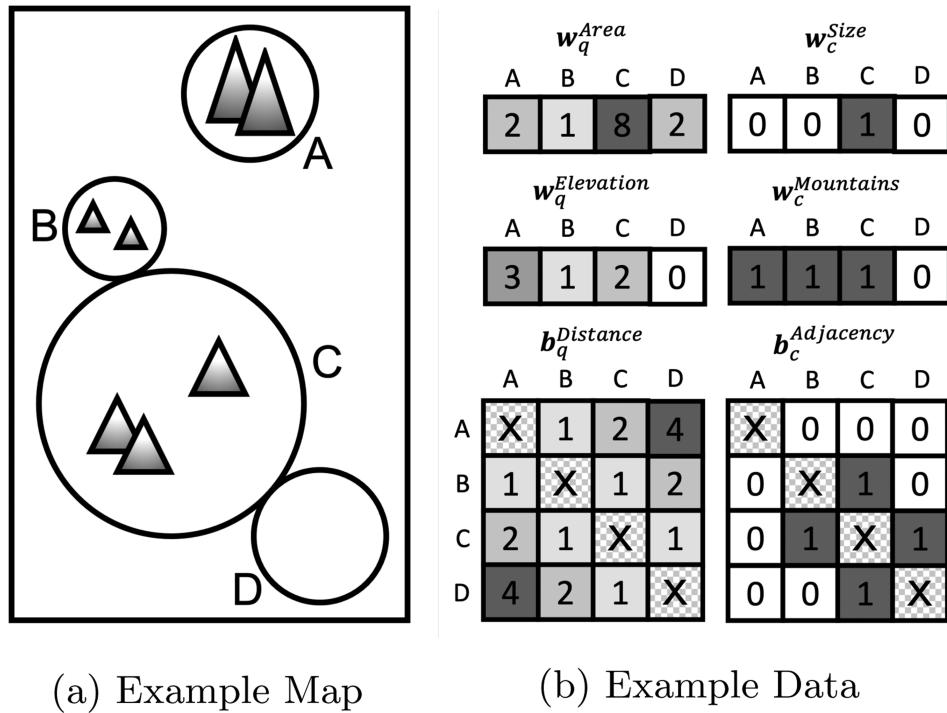


FIGURE 2. (a) A visual representation of a hypothetical biogeographic system with four regions (A, B, C, D) with different quantitative (area, elevation, distance) and categorical features (size, mountain-presence, adjacency). (b) Regional features of the biogeographic system are encoded by MultiFIG into variables containing quantitative data (numerical values in w_q and b_q), categorical data (binary values in w_c and b_c), within-region data (vectors w_c and w_q), and between-region data (matrices b_c and b_q).

but are ultimately transformed so that they can collectively contribute to evolutionary rates. The MultiFIG model also separates data by dimensionality type. Data can represent 1D within-region information (values for each individual region) or 2D between-region information (values for each region pair). We use four different containers to represent these data types. Each of these containers can store data for several different features layers of a given type that are indexed by k . An example of a hypothetical four-island system with six regional features of different types can be seen in Figure 2.

Quantitative feature data are stored in the N -vector w_q^k for within-region features and in the $N \times N$ -matrix b_q^k for between-region features, where N is the number of regions in the system. For instance, w_q^1 might contain areas (e.g., km^2) for each region, while w_q^2 may contain mean elevations for each region. Vectors in w_c and matrices in b_c are used similarly to represent data about within- and between-region categorical data. Note that the feature layer indexed by a specific value of k in one container does not necessarily imply anything about the layer indexed by the same value of k in another container—e.g., layers indexed by $k = 2$ do not all necessarily describe features associated with elevation. For instance, w_c^1 might contain a vector categorizing

regions as “small/large,” and b_c^1 might contain a matrix representing region pair adjacency/nonadjacency. The processes of extinction and within-region speciation make use of within-region data, while the processes of dispersal and between-region speciation make use of between-region data. Table 1 describes the different features used in our empirical analysis, giving an example of the types of data that may be included in these different containers. However, the framework is general, and features can be added or substituted as needed for different study systems.

The MultiFIG model associates regional feature data with biogeographic rates through a series of parameterized transformations (Fig. 3). The function q_x^k is used to handle quantitative data and parameterizes the effects of a quantitative feature layer k on the biogeographic process x , where x can represent dispersal (d), extinction (e), within-region speciation (w), or between-region speciation (b). Each quantitative feature layer k is assigned a single set of ϕ_x^k parameters. These parameters represent the strength and direction of the effect of each feature on a particular process (ϕ_e^k for extinction, ϕ_w^k for within-region speciation, ϕ_d^k for dispersal, and ϕ_b^k for between-region speciation). All regions or region pairs share one ϕ_x^k per feature and process. For each feature layer k and region i (or region pair i, j), the value of the feature for the region/pair is normalized by the geometric mean of

TABLE 1. Overview of features and MultiFIG parameters.

Feature	Description	Processes	Parameters
w_q^{Area}	Areas of each region	w, e	$\phi_w^{Area}, \phi_e^{Area}$
$w_q^{Elevation}$	Mean elevations of each region	w, e	$\phi_w^{Elevation}, \phi_e^{Elevation}$
$b_q^{Distance}$	Mean distances between each region pair	b, d	$\phi_b^{Distance}, \phi_d^{Distance}$
$b_q^{Elevation}$	Differences in mean elevation between each region pair	b, d	$\phi_b^{Elevation}, \phi_d^{Elevation}$
w_c^{Area}	Sizes of each region (large/small)	w, e	$\sigma_w^{Area}, \sigma_e^{Area}$
$w_c^{Elevation}$	Heights of each region (highland/lowland)	w, e	$\sigma_w^{Elevation}, \sigma_e^{Elevation}$
$b_c^{Distance}$	Adjacency of each region pair (adjacent/nonadjacent)	b, d	$\sigma_b^{Distance}, \sigma_d^{Distance}$
$b_c^{Elevation}$	Comparison of heights of each region pair (same/different)	b, d	$\sigma_b^{Elevation}, \sigma_d^{Elevation}$

Note: Each feature (**w** or **b**) is associated with two of four processes: within-region speciation (w), extinction (e), between-region speciation (b), and dispersal (d). The strength of the effect is controlled through a set of estimated parameters (ϕ or σ). Details provided in the main text.

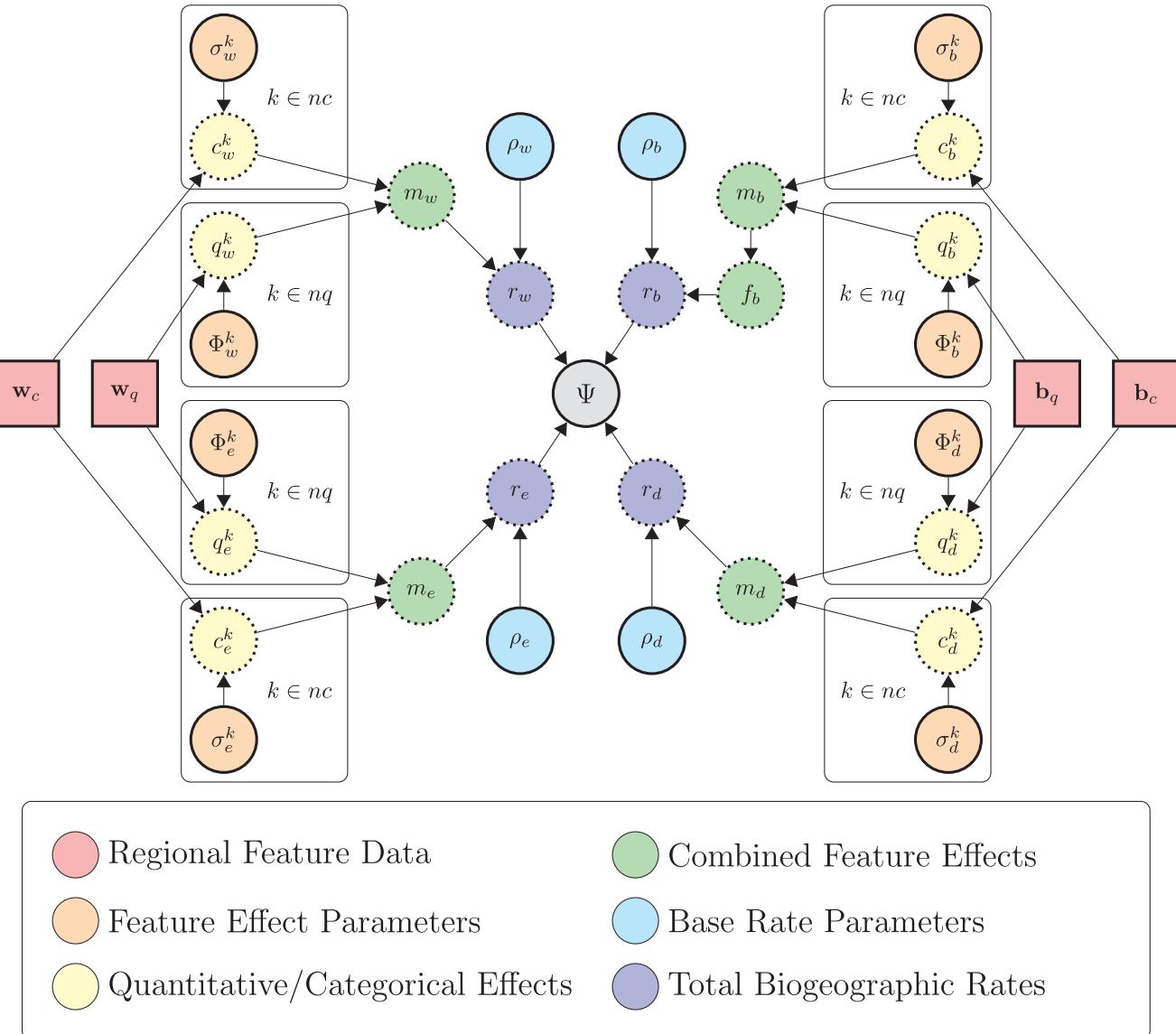


FIGURE 3. Full graphical model for generalized MultiFIG. The main text and table describe how to interpret regional feature data (**w**, **b**; red), feature effect parameters (ϕ , σ ; orange), individual feature effects (q , c ; yellow), combined feature effects (m) and the split score function (f_b ; green), base rates (ρ ; cyan), the absolute biogeographic rates (r ; blue), and the GeosSE process (Ψ ; gray). Graphical model notation follows the conventions described in Höhna et al. (2014). Square nodes represent constant values (data). Circle nodes with solid lines represent stochastic variables (model parameters, and the phylogeny, which is fixed in this analysis). Circle nodes with dotted lines represent deterministic variables (functions). Large rectangles indicate iterative plates.

the feature over all regions/pairs, $gm(\cdot)$ and then exponentiated by the effect parameter ϕ_x^k . Thus, the vector q_x^k represents the effects of each quantitative feature k on a particular process x and contains information about each region i or region pair i,j .

$$q_x^k(i) = \left[\frac{\mathbf{w}_q^k(i)}{gm(\mathbf{w}_q^k)} \right]^{\phi_x^k} \quad (1)$$

$$q_x^k(i,j) = \left[\frac{\mathbf{b}_q^k(i,j)}{gm(\mathbf{b}_q^k)} \right]^{\phi_x^k}. \quad (2)$$

Similarly, the function c_x^k is used to handle categorical data and parameterizes the effects of a categorical feature layer k on process x . Each categorical feature layer k is assigned a single set of σ_x^k parameters. These parameters represent the strength and direction of the effect of each feature on a particular process (σ_e^k for extinction, σ_w^k for within-region speciation, σ_d^k for dispersal, and σ_b^k for between-region speciation). All regions or region pairs share one σ_x^k per feature and process. For each feature layer k and region i (or region pair i,j), an indicator function is used to determine which category the region/pair falls into (here labeled A and B). If the region/pair is in category A , the indicator function produces 0, and if it is in state B , the indicator function produces 1. This indicator function is used to exponentiate the $\exp(\sigma_x^k)$ parameter, which means the c_x^k function will output 1 if the region/pair is in category A , and $\exp(\sigma_x^k)$ if the region/pair is in category B . Thus, the vector c_x^k represents the effects of each categorical feature k on a particular process x and contains information about each region i or region pair i,j .

$$\mathbb{1}(i,k) = \begin{cases} 0 & \mathbf{w}_c^k(i) = A \\ 1 & \mathbf{w}_c^k(i) = B \end{cases} \quad (3)$$

$$\mathbb{1}(i,j,k) = \begin{cases} 0 & \mathbf{b}_c^k(i,j) = A \\ 1 & \mathbf{b}_c^k(i,j) = B \end{cases} \quad (4)$$

$$c_x^k(i) = [\exp(\mathbb{1}(i,k) \times \sigma_x^k)] \quad (5)$$

$$c_x^k(i,j) = [\exp(\mathbb{1}(i,j,k) \times \sigma_x^k)]. \quad (6)$$

The form of this equation (making use of an indicator function) is more general than necessary for the analyses presented here, since we represented all categorical data according to a 0/1 binary. However, it displays how the framework could be adapted for more complicated sets of possible categorical values.

The m_x function combines the effects of all categorical and quantitative feature layers for a given region i or region pair i,j . Each process has its own m_x vector containing information about the total effects of all features on each region or region pair ($m_e(i)$, $m_w(i)$, $m_d(i,j)$, $m_b(i,j)$). Note that parameter values of $\phi_x^k = 0$ or $\sigma_x^k = 0$ yield $q_x^k = 1$ and $c_x^k = 1$ for all regions i or region pairs i,j , indicating that the feature will not influence region-specific event rates (a null effect).

Relative rate of extinction or within-region speciation in region i is computed as the product over all relative rate contributions for all relevant quantitative and categorical effects for region i , with the general form

$$m_x(i) = \prod_k c_x^k(i) \times \prod_k q_x^k(i), \quad (7)$$

whereas analogous dispersal and between-region speciation rate modifiers involving regions i and j have the form

$$m_x(i,j) = \prod_k c_x^k(i,j) \times \prod_k q_x^k(i,j). \quad (8)$$

Finally, the m_x values are scaled by a base rate parameter for each process, ρ_x . As before, the absolute extinction or within-region speciation rate for region i is computed as

$$r_x(i) = \rho_x \times m_x(i), \quad (9)$$

the absolute dispersal rate between regions i and j is computed as

$$r_d(i,j) = \rho_d \times m_d(i,j), \quad (10)$$

while the absolute between-region speciation rate for splitting range R into daughter ranges S and T is

$$r_b = \rho_b \times f_b(S, T; m_b), \quad (11)$$

where f_b is the range split score function, which is used to calculate the relative rates of between-region speciation for different splits (Landis et al. 2022), and is defined as

$$f_b(S, T; m_b) = \frac{1}{gm(f_b)} \times \left[\sum_{(i,j) \in E(S,T)} \left[\frac{m_b(i,j) + m_b(j,i)}{2} \right]^{-1} \right]^{-1}. \quad (12)$$

In effect, f_b assigns smaller values to ranges that are more "cohesive," where cohesiveness is measured by how easily the current range $R = S \cup T$ can be split into subranges S and T given the strength of connectivity between regions modeled by m_b . Let the original range be a fully connected graph, while treating occupied regions in $R = S \cup T$ as nodes and the matrix m_b as edge weights.

$E(S, T)$ represents the cutset, which is the minimal set of edges that must be cut to fully separate the nodes in range S from the nodes in range T .

This produces a set of absolute process rates $r_e(i)$, $r_w(i)$, $r_d(i, j)$, and $r_b(S, T)$ for all regions in the biogeographic system, representing the rates of extinction, within-region speciation, dispersal, and between-region speciation in the GeoSSE model. However, unlike in the GeoSSE framework, every distinct biogeographic event is not assigned its own free rate parameter; they are instead modeled using observable data and a much smaller set of parameters that remains constant with regard to the number of regions in the system (Fig. 3).

Here, we have presented a model where feature effects are all independently multiplicative. This simple model will be used throughout our simulation study and empirical analysis to demonstrate MultiFIG's usefulness for identifying individual features of importance. However, for the investigation of more detailed hypotheses, the c , q , and m functions can be modified to produce different relationships among features and effect parameters. For example, the MultiFIG framework can be used to reproduce the original FIG model, which was designed to test hypotheses about *Lioleumus* lizards. In this model, each categorical feature effect modifies one quantitative feature effect, which led to the finding that distance matters for dispersal—but over-water distance matters more than over-land distance (Landis et al. 2022). One could even explore more complex models with additional parameters specific to different regions or pairs and even perform model selection among these models with RJMCMC. However, incorporating too many parameters would cause the model to encounter the same computational difficulties as a many-region GeoSSE model and may confound the broad-scale relationships that MultiFIG was designed to investigate.

Model Selection

The MultiFIG modeling framework allows for the inclusion of numerous regional features that may or may not correlate with evolutionary rates. Typically, determining which parameters are or are not necessary in a model requires one to test each possible combination of parameters and choose the best model according to some criteria. For example, one could use Bayes factors computed using marginal likelihoods to compare models, requiring independent MCMC runs for each model being compared. However, using reversible-jump Markov chain Monte Carlo (RJMCMC), it is possible to "jump" between models with different dimensionality (Green 2003). Because of this, RJMCMC allows us to perform model selection in addition to parameter estimation. In practice, a new proposal is added to the MCMC for each feature effect parameter ϕ or σ , which suggests that the parameter's value is actually equal to 0, effectively turning the parameter "off." The reverse proposal is also added,

turning the parameter back "on" and allowing the MCMC to continue to explore nonzero parameter values. For simple nested models, the Metropolis–Hastings–Green ratio for transdimensional proposals is very straightforward to compute, assuming that the priors distributing reversible-jump variables assume independence (Green 1995). RevBayes allows for the implementation of reversible-jump MCMC moves (Höhna et al. 2019). We can keep track of the amount of time the Markov chain spends with a parameter "on" versus "off," which can be used to obtain a posterior probability. In this way, we are able to examine whether various features have an effect on a particular process, or whether they have no effect (an effect parameter value of 0). Note that the term "effect" here refers to a mathematical relationship and does not necessarily indicate causality; a feature with an inferred effect may not have contributed to diversification or biogeographic rates directly. Rather, it may be correlated with some other, unexamined feature that influenced these processes.

Since geographical features themselves are often correlated, especially when both quantitative and categorical versions of the same feature are incorporated, it is essential that we are able to interpret reversible-jump posteriors with respect to the "on/off" status of other parameters. Fortunately, RJMCMC results allow us to examine the joint probabilities of different parameter combinations. For example, we might obtain the following diagram representing joint probabilities for the effect/noneffect of two parameters connected to closely related phenomena, such as highland/lowland (σ) and mean elevation (ϕ).

$$\begin{array}{ccc} & \sigma = 0 & \sigma = 1 \\ \phi = 0 & 0.01 & 0.49 \\ \phi = 1 & 0.49 & 0.01 \\ \text{marginal} & 0.50 & 0.50 \end{array} \quad \left(\begin{array}{ccc} & \sigma = 0 & \sigma = 1 & \text{marginal} \\ \phi = 0 & 0.01 & 0.49 & 0.50 \\ \phi = 1 & 0.49 & 0.01 & 0.50 \\ \text{marginal} & 0.50 & 0.50 & - \end{array} \right).$$

Marginal probabilities represent the probability of an event independent of other events and are typically returned by MCMC analyses. By examining the marginal probabilities associated with each parameter in the above example, it could not be concluded that either parameter has an effect. For instance, the marginal probability of $\phi = 1$ is 0.50, and the marginal probability of $\phi = 0$ is 0.50. However, by examining the joint probabilities, we find that the scenarios where one of the parameters has an effect are much more probable than the scenario where neither parameter has an effect. The probability of $\phi = 1, \sigma = 0$ is 0.49, the probability of $\phi = 0, \sigma = 1$ is 0.49, and the probability of $\phi = 1, \sigma = 1$ is 0.01, for a total probability of 0.99. The probability of $\phi = 0, \sigma = 0$ is 0.01. Therefore, we could conclude that elevation, either in a categorical or quantitative context, likely has an effect on a certain process, even if the specifics (categorical or quantitative) are confounded by the model design.

Coverage Experiment

To assess the validity and performance of our model when correctly specified, we simulated 900 biogeographic data sets under a range of scenarios that varied in terms of the number of simulated taxa, the number of regions, the values of regional features, and the values of evolutionary process parameters.

Using R, we simulated 900 different geographic scenarios with 3, 5, or 7 regions (150, 600, and 150 geographies, respectively). Each region/region pair within each geographic scenario was assigned a value for eight different region features. Four were quantitative: area, mean elevation, difference in mean elevation, and distance. The other four were categorical representations of the same quantitative features. Since all region features would later be normalized, the scales of the features were not selected to match a particular biological scenario. However, the variances of the features were designed such that the *Liolaeus* study system of [Esquerre et al. \(2019\)](#) would fall comfortably within the range of simulation scenarios tested (Fig. 4).

Region areas were drawn randomly from a gamma distribution with shape 1 and rate 0.15. Categorical areas were then assigned to regions based on whether a region's area was greater than (1) or less than (0) the mean area of all regions in that geographic scenario.

Mean elevations were drawn from the same gamma distribution, and categorical elevations were assigned in the same manner. Quantitative differences in elevation were calculated by taking pairwise differences between each region. Categorical differences in elevation were assigned based on whether each pair of regions had the same categorical state (1) or different states (0). To preserve distance properties, each region was assigned a "location" consisting of x and y coordinates drawn from a uniform distribution on $[0, 1]$. Euclidean distances were then calculated between each region pair. Categorical distances (adjacency) were not calculated based on quantitative distances, as this does not always represent physical reality (as in the *Liolaeus* study system). Rather, region pairs were randomly assigned adjacency (1) or nonadjacency (0) with equal probability.

From the simulated geographies, we used RevBayes ([Höhna et al. 2016; Freyman and Höhna 2017](#)) to generate 900 simulated time-calibrated phylogenies and biogeographies. Using the 600 geographies with 5 regions, we created 600 phylogenies/biogeographies, with 150 simulations targeted at each tree size category: extra small (25–50 taxa), small (51–100 taxa), medium (101–200 taxa), and large (201–350 taxa). This was done through rejection sampling; if the model parameters

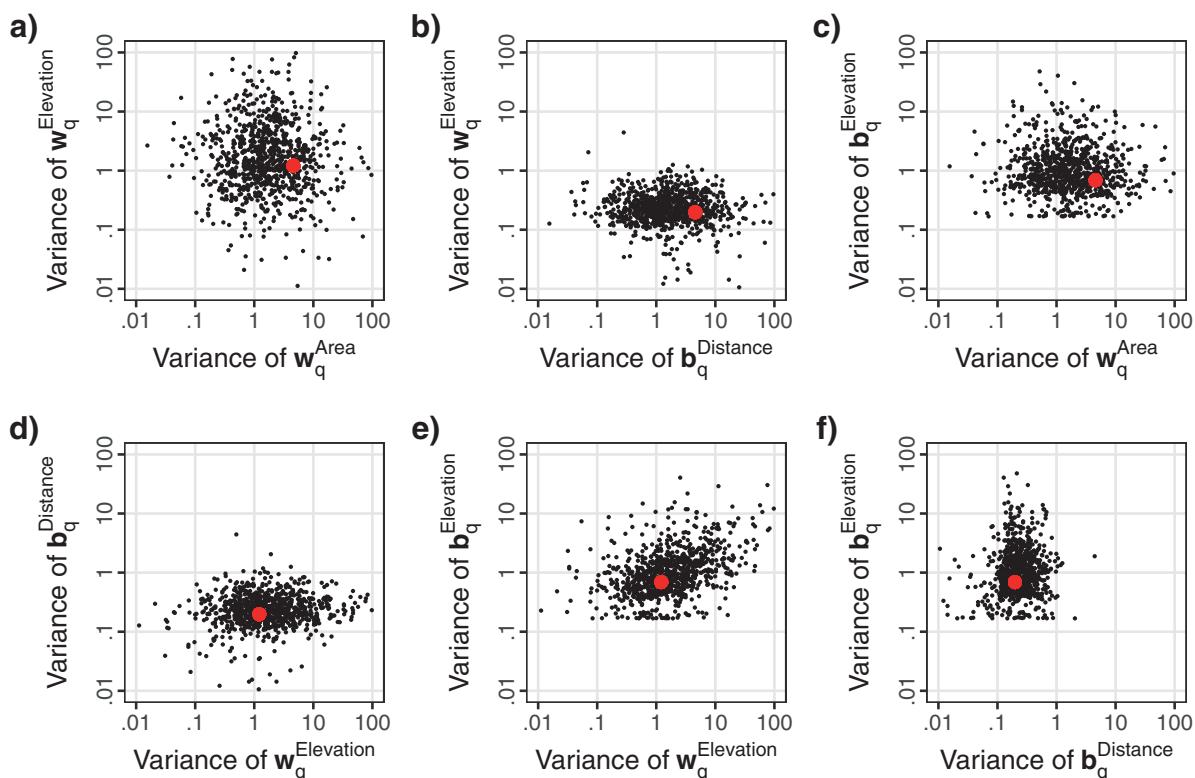


FIGURE 4. Pairwise variances of geographic features for 900 simulated geographies. Variances of regional features among the six South American regions used in our *Liolaeus* MultiFIG analysis are also included (red dots). Simulated scenarios cover a wide range of feature variances, and the variances of the South American regional features are in well-explored regions of parameter space.

randomly drawn by RevBayes did not create an appropriate tree, we redrew new parameters and attempted to simulate a new tree. First, we checked whether the tree was in the appropriate size class, and rejected the tree if it had too many or too few tips. Then, we tested whether the tree would generate a valid likelihood, and rejected the tree if it did not. Typically, trees rejected for invalid likelihoods had near-concurrent events (within an extremely short [1e-9] interval of time)—these trees are sometimes generated during simulation due to conditioning or rounding but are assigned zero probability by TensorPhylo because the MultiFIG model does not contain events that occur simultaneously (e.g., mass speciation events). We repeated this process until one adequate tree was generated for each geography. Simulations were conditioned on time, so all resulting trees had a height of 1. The other 300 geographies (150 with 3 regions and 150 with 7 regions) were used to simulate 300 phylogenies/biogeographies with medium-sized trees using the same process.

Simulations were performed under the MultiFIG model using RevBayes. In simulations used for coverage testing, all within-region features were modeled to act upon both within-region processes (extirpation and within-region speciation), whereas all between-region features were modeled to act upon both between-region processes (dispersal and between-region speciation). Feature effect parameters were drawn from a normal distribution with mean 0 and variance 1, while base rate parameters for each process were drawn from an exponential distribution with a rate of 1. A reversible-jump model was not used during these simulations (no feature effect parameters were given a zero value) because RJMCMC creates point masses in posteriors, which would complicate the interpretation of HPD intervals for the coverage experiment. A second set of simulations with various features turned “on” or “off” was used to investigate different types of model misspecification.

The simulations resulted in 900 unique phylogenies and sets of tip states (ranges of extant taxa). Using RevBayes (Höhna et al. 2016) and Tensorphylo (May and Meyer 2022), these simulated phylogenies/biogeographies were then analyzed under the generative model using standard (not reversible-jump) MCMC, to determine model performance in different regions of parameter space. MCMC analyses were conducted over 10,000 generations, discarding 10% as burn-in (without tuning hyperparameters). Of these analyses, 787 ran to completion; the other analyses either failed in low-likelihood regions of parameter space or did not complete 10,000 generations within the 2 week runtime limit imposed on jobs by our cluster.

Using R, we summarized the Bayesian coverage of each model parameter, including base rates ρ for each process and feature-associated σ and ϕ parameters. For the simulation study only, we utilized 80% highest posterior density (HPD) intervals to calculate coverages. We chose to use HPD80 because HPD95 is expensive to estimate accurately, and the goal of the coverage

analysis is to determine model performance, not to infer significance of particular parameters. If the inference machinery is behaving properly and the model is correctly specified, we would expect a coverage value of 80% (80% of 80% HPD intervals should cover the truth). We also calculated median absolute percent error for different scenarios, indicating the expected accuracy of analyses fitting certain conditions.

Model Misspecification Experiment

To assess how the MultiFIG model performs under various types of model misspecification, we performed a set of 200 additional simulations under 4 conditions (50 simulations per condition) using reversible-jump MCMC. In these simulations, feature effect parameter values were not drawn from a distribution; instead, all “on” features were assigned an effect parameter with value 1. This was done to control for the magnitude of effect when assessing MultiFIG’s ability to correctly identify “on” features across conditions. We used the same rejection sampling procedure to simulate trees as in the coverage experiment. All simulations were analyzed under the same model using the 8 regional features described above: categorical and quantitative area, categorical and quantitative distance, categorical and quantitative within-region elevation, and categorical and quantitative between-region elevation. Reversible-jump moves assigned a probability of 0.5 to each parameter being “on.”

Under the first condition, labeled “correct,” we simulated trees using all 8 regional features and analyzed the data sets with those same 8 features. This was done to examine the power of RJMCMC to identify “true” simulating parameters when the model is correctly specified. Of these analyses, 46 ran to completion.

Under the second condition, labeled “extra feature,” we simulated trees using only 4 regional features: categorical and quantitative area, and categorical and quantitative distance. The remaining 4 features were not used to inform rates of diversification or dispersal. We then analyzed these data sets under the full 8-feature model to assess whether MultiFIG would distinguish between the “true” features which were used in simulation and the “false” features which were not. These simulations were also used to assess whether including fewer simulating features would improve MultiFIG’s ability to identify those features. Of these analyses, 41 ran to completion.

Under the third condition, labeled “missing feature,” we simulated trees using 12 regional features, including the 8 regional features used in previous analyses and an extra set of 4 features arbitrarily called “temperature”: categorical and quantitative within-region temperature, and categorical and quantitative between-region temperature. These features were drawn in the same manner as the “elevation” feature. We analyzed these data sets under the 8-feature model to determine whether MultiFIG would correctly identify “true” features

even when other unidentified features also informed the true process rates. Of these analyses, 43 ran to completion.

Under the fourth condition, labeled "noisy," we simulated trees using 8 regional features. Four of these features (categorical and quantitative area, and categorical and quantitative distance) were drawn in the same manner as above. We also simulated categorical and quantitative within-region elevations, and categorical and quantitative between-region elevations. However, before using these 4 features to generate a tree, we added noise to these parameters by drawing new values for each region/pair from a uniform distribution (lower = $0.5 \times$ original value, upper = $1.5 \times$ original value). This new feature set was labeled "true elevation" because it was used in simulation; the other elevation feature set was labeled "noisy" and was used during analysis. We analyzed these data sets under the 8-feature model to assess whether MultiFIG would still assign significance to the "noisy" features and whether the addition of noise would alter MultiFIG's ability to identify "true" features. Of these analyses, 45 ran to completion.

Liolaemus Analysis

We applied our FIG model to an empirical study system, the *Liolaemus* lizards of South America, in order to demonstrate how our framework can reveal relationships between region features and evolutionary rates. We split the southern part of the continent of South America into 6 discrete regions for the MultiFIG analysis, as outlined in Esquerre et al. (2019), and generated spatial polygon representations of these regions using QGIS. Elevation data were obtained from the NASA Shuttle Radar Topographic Mission (NASA 2013) using the "raster" package in R. We extracted data for 8 regional features (Table 1): quantitative area (W_q^{Area}), categorical area (above/below mean, W_c^{Area}), quantitative distance (mean distance between a randomly selected point in the first region and a randomly selected point in the second region, B_q^{Distance}), categorical distance (adjacent/nonadjacent, B_c^{Distance}), quantitative elevation (mean elevation, $W_q^{\text{Elevation}}$), categorical elevation (above/below mean, $W_c^{\text{Elevation}}$), quantitative elevation difference (difference in mean elevation, $B_q^{\text{Elevation}}$), and categorical elevation difference (whether regions share highland/lowland status or not, $B_c^{\text{Elevation}}$).

The time-calibrated phylogeny of *Liolaemus* and present-day ranges were obtained from Esquerre et al. (2019). While the Esquerre et al. data set includes data for the entire Liolaemidae family, we focus on the genus *Liolaemus*. The other two genera are older, less species-rich, and more narrowly distributed than *Liolaemus*. *Ctenoblepharys* is monotypic, and *Phymaturus* predominantly consists of Patagonian and Central Andean species. Therefore, to avoid the potential biases that these genera could introduce because of

model misspecification, we exclude *Ctenoblepharys* and *Phymaturus*. We fixed the phylogeny in this analysis due to data availability. To infer rates of evolutionary processes and effect parameters associated with each feature, we used RevBayes (Höhna et al. 2016) to perform an RJMCMC analysis. Due to time constraints and computational challenges, two separate chains were run concurrently and concatenated (after removing 10% burn-in from each chain, which was used to tune hyperparameters). Several independent runs were compared to assess consistency, but the results presented here are drawn from one concatenated RJMCMC analysis, as all analyses showed the same patterns. Priors were identical to those in the simulation experiments, except the normal distributions on feature effect parameters were widened to have a standard deviation (SD) of 2.

We analyzed the results using R and determined which σ and ϕ parameters indicated strong effects (significant probability of nonzero value for reversible jump indicator or zero value outside 95% HPD interval). We also estimated base rates of each process (Supplementary Fig. S9) and generated an ancestral state tree using the MCMC output (Supplementary Fig. S14).

MultiFIG Code and Tutorial

Raw data, R scripts, and RevBayes scripts used in these analyses are available in the code repository <https://bitbucket.org/sswiston/multifig/src/master/> and the Dryad Digital Repository <https://doi.org/10.5061/dryad.dz08kps2x>. The preconfigured Docker container used to run RevBayes MultiFIG scripts is available on Docker Hub: https://hub.docker.com/r/sswiston/rb_tp (image 7). The RevBayes website hosts a tutorial describing how to design a MultiFIG analysis for other clades and biogeographic systems: <https://revbayes.github.io/tutorials/multifig/>.

RESULTS

Simulation Experiment

Coverages were calculated for each model parameter, representing the number of data sets where the 80% HPD interval of the estimate contains the "true" value of the parameter (the value used in simulation). The coverage analysis indicates that the inference machinery behaves as expected (Supplementary Table S3). However, mean coverage (77.8%) is slightly below the 80% benchmark. This could possibly be due to the way in which trees were generated. By targeting specific tree size categories, many trees were rejected. Also, if the parameter values drawn from the prior were unable to create a valid tree, new parameter values had to be selected. Therefore, the true distribution of simulating parameters may not precisely match the priors used for simulation, causing a slight decrease in coverage due to misspecification. This does not indicate a problem with the model itself.

MultiFIG estimates evolutionary rates and feature effect parameters with a reasonable degree of accuracy (Fig. 5). Median absolute percent errors give a more detailed view of the model's overall performance (Supplementary Fig. S8). Generally, the model produced more accurate estimates of base rates in scenarios with larger trees, and it produced more accurate estimates of effect parameters (ϕ and σ) in scenarios with more regions.

Results of reversible-jump simulations (Table 2) indicate negligible rates of mistakenly assigning probability >0.95 to a noneffect for a "true" feature used in simulation or assigning probability >0.95 to an effect for a "false" feature not used in simulation. When the level of significance was decreased to 80%, the rate of incorrectly selecting "false" features remained low at 0.015, while the rate of correctly selecting "true" features was

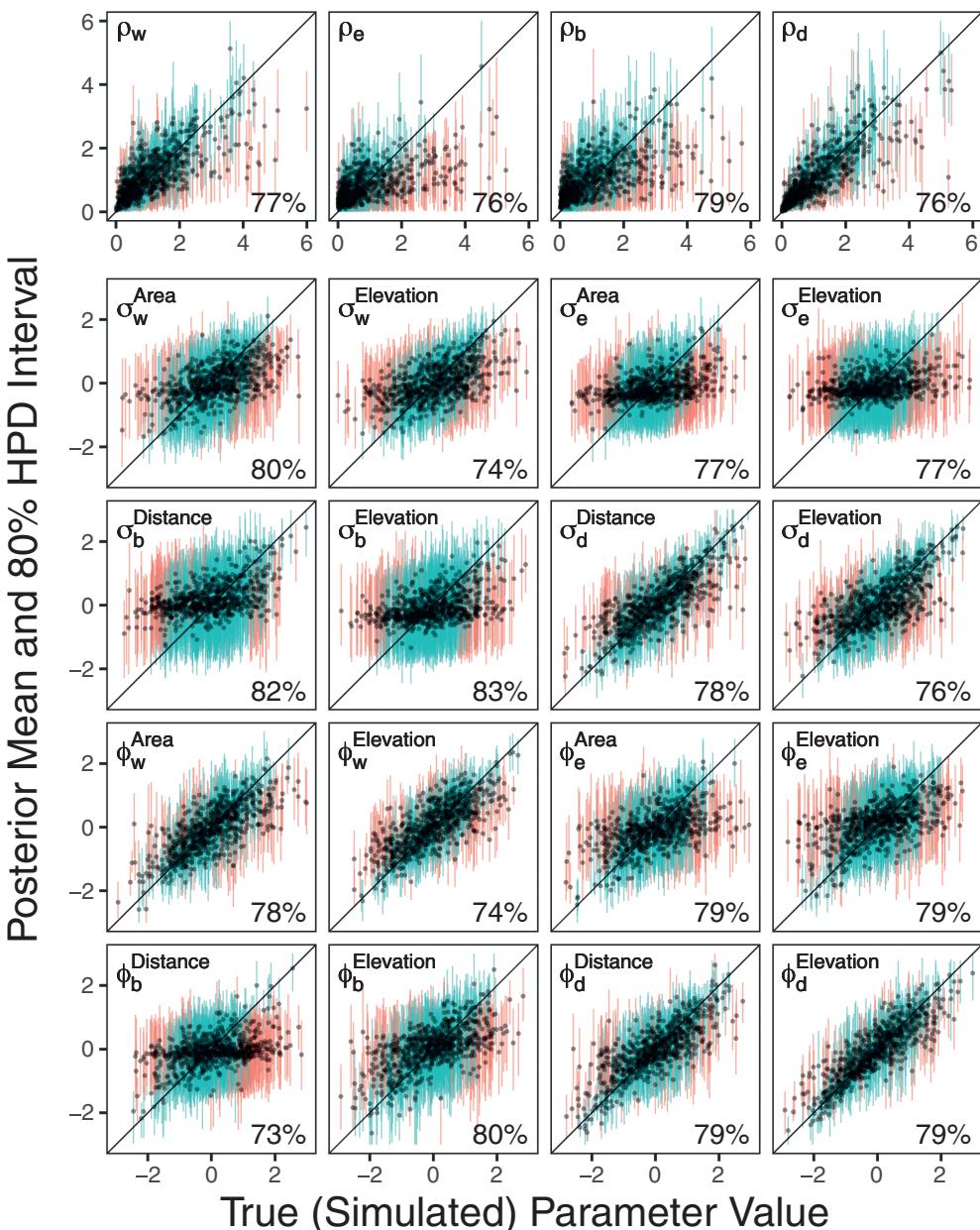


FIGURE 5. Quality of estimation and coverage of MultiFIG model parameters in simulations with five regions and medium tree size. Plots show true parameters on the x -axis and estimated values on the y -axis. Parameters are given in the upper left, and coverage percentages in the lower right of each plot, and 80% HPD intervals, which cover the truth, are shown in blue, and 80% HPD intervals, which do not cover the truth, are shown in red. Some HPD intervals were clipped to the bounds of the plot (0 to 6 for ρ parameters, -3 to 3 for σ and ϕ parameters). Coverage analysis indicates that the inference method behaves appropriately, and plots indicate that most parameters can be inferred reasonably well (demonstrated by alignment with the 1:1 line).

TABLE 2. True and false positive rates for model misspecification experiment.

Model	Parameters	Probability of 80% Significance
Correct	True	0.196
	False	0.015
Extra feature	True	0.271
	Noisy	0.208
Missing feature	True	0.214
	Noisy	0.2

an order of magnitude greater under all conditions. This implies that, if MultiFIG identifies a feature during an analysis, it is very likely to be a "true" feature (i.e., actually correlated with a process). Importantly, failing to include all simulating features ("missing feature" condition) or adding noise to some parameters ("noisy" condition) did not substantially alter MultiFIG's ability to correctly identify any features, whether "true" or "noisy." In an 8-feature model, MultiFIG was able to correctly select "true" features at a rate of ~0.2, making the model very conservative. As expected, a simpler generating model ("extra feature" condition with fewer simulating features) resulted in more power for identifying "true" parameters.

Lioleamus Analysis

Here, we present posterior intervals alongside reversible-jump probabilities for each parameter/process pairing (Supplementary Figs S10–S13). Note that the posterior densities have a point mass at 0 due to the reversible jump component of the analysis. Because the HPD intervals generally assume unimodal densities without point masses, we will rely primarily on reversible jump probabilities for assigning significance to feature effect parameters to simplify the presentation. We also present joint RJMCMC probabilities for quantitative and categorical representations of the same features (Fig. 6).

Of the 16 feature effect parameters examined, 4 were inferred to have a probable effect, uncovering several interesting relationships between features and processes. First, posterior densities indicate a probable negative relationship between geographical distance and dispersal. The results imply that both quantitative distance (mean distance between points in each region) and categorical distance (region adjacency) are important components of the model. Second, an analysis of joint RJMCMC probabilities of highly correlated effect parameters (categorical and quantitative representations of the same feature) indicate that region area (size) is negatively correlated with rates of extinction. The model assigns a low probability (0.03) to the scenario where neither the quantitative nor categorical parameter has an effect, meaning that region area likely relates to extinction in some capacity. However, the model does not assign a high probability of both parameters being simultaneously included (0.28) and does not seem to

have a strong preference for either the quantitative or categorical representation.

The remaining 12 feature effect parameters were not inferred to have an effect. In these instances, RJMCMC results were unable to reject the null (noneffect), and joint RJMCMC probabilities did not indicate the importance of either the quantitative or categorical representation of a particular feature. The model was also unable to strongly reject most parameters, with the exception of ϕ_w^{Area} and $\phi_w^{\text{Elevation}}$. Interestingly, one of the feature/process pairings where we did not find an effect was elevation on within-region speciation. The parameters $\phi_w^{\text{Elevation}}$ and $\sigma_w^{\text{Elevation}}$ represent the impact of quantitative and categorical elevation on within-region speciation rates. The null effect (0) is well within the 95% posterior interval for both of the parameters, and the reversible jump MCMC supports a 0 value for both. Additionally, the joint RJMCMC analysis assigns highest probability to the scenario where neither parameter is incorporated ($P = 0.83$). Therefore, our model did not find any significant effect of quantitative or categorical elevation on within-region speciation.

Examining the posterior distribution for the number of "on" parameters gives a more nuanced picture (Fig. 7). Here, we can see overrepresentation of three of the four parameters ($\sigma_d^{\text{Distance}}$, ϕ_d^{Distance} , and σ_e^{Area}) preferred by the individual/pairwise analyses. We can also see underrepresentation of some parameters (such as ϕ_w^{Area} and $\phi_w^{\text{Elevation}}$), which indicates that the model prefers not to include these parameters. The model prefers scenarios with more than four parameters but does not strongly suggest which parameters may be involved.

DISCUSSION

Our study explored how MultiFIG can be used to test and measure the many ways that regional features might shape corresponding biogeographic rates of dispersal, speciation, and extinction. Through simulation, we demonstrated the ability for MultiFIG to infer the effects of regional features on evolutionary processes. For the number of features examined (eight quantitative and eight categorical), the model performed best in scenarios with more regions and larger trees (Landis et al. 2022). This is a sensible outcome, as we would expect an increase in the amount of data (more known feature values and more observed taxa) to improve parameter estimates. In all scenarios, the coverage analysis indicates that MCMC results can be trusted; if the 95% HPD interval "rules out" a null effect, we can be reasonably confident that a feature is related to an evolutionary process. From the examination of RJMCMC analyses, we can see that the reversible-jump model component has limited power to detect effects, particularly in complex systems with many true feature effects. However, the rate of significant false negatives was negligible, and the

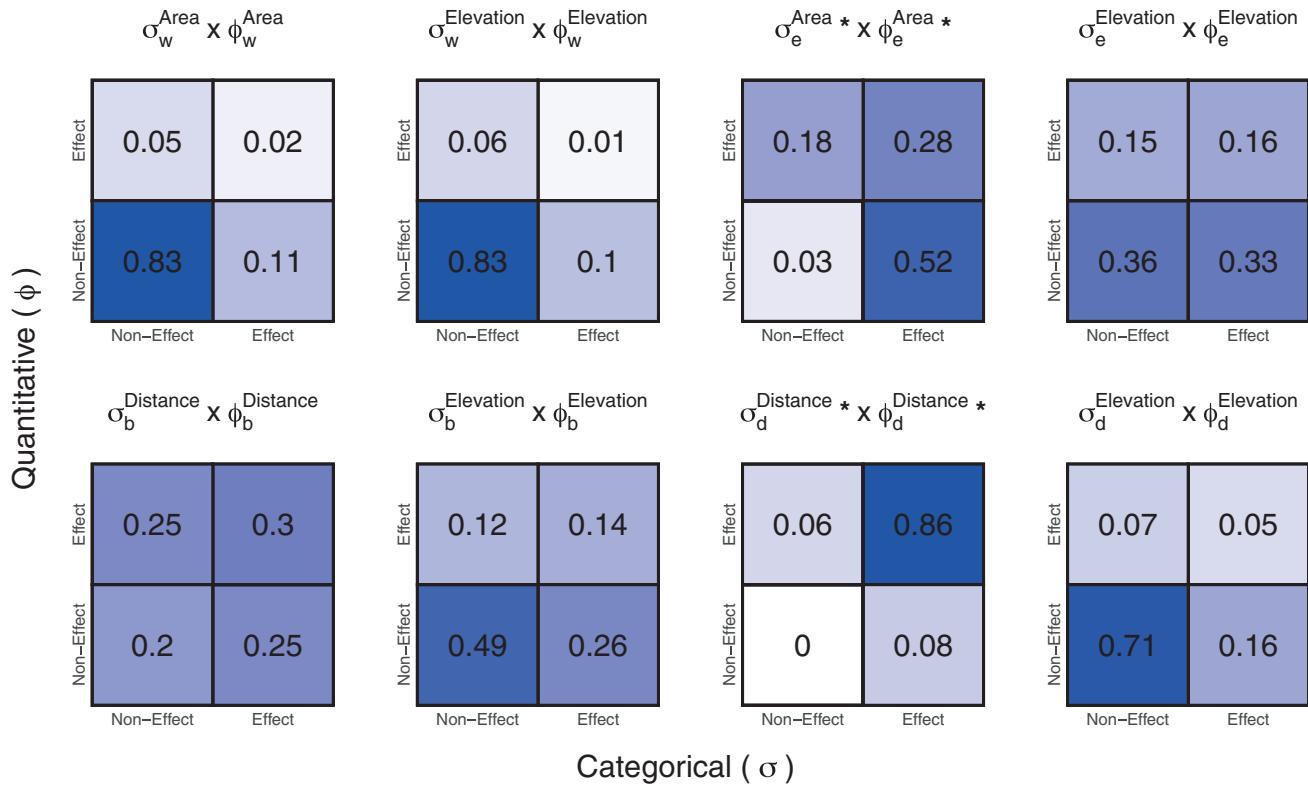


FIGURE 6. Joint probabilities of highly correlated MultiFIG model parameters in the reversible-jump analysis of *Liolaemus*. The lower-left quadrant indicates the probability that a feature does not affect a certain process, since neither parameter associated with the feature (quantitative or categorical) shows an effect. The only feature/process pairings for which we can reject the null model (noneffect) are elevation/dispersal and area/extinction.

rate of significant false positives was very low, even in the presence of extra features, missing features, and feature noise. Therefore, if RJMCMC produces a significant probability of either effect or noneffect, there is good reason to trust that result. This can be useful in conjunction with traditional posterior estimates and HPD intervals.

Overall, the simulation study shows that MultiFIG has the ability to infer relationships between features and processes under a correctly specified model and under several common types of model misspecification. If we believe we have incorporated some of the features which are most closely related to evolutionary processes, and accept the assumption that all features act independently on evolutionary rates, we can be reasonably confident in our results. This remains true even when we fail to capture all relevant features, include unrelated features, or select “noisy” features (e.g., features with measurement error, or features that are correlated proxies for “true” features). Other types of model misspecification remain to be studied. For instance, there may be interactions between features that we fail to capture, which can be modeled by treating those features as structured random variables (e.g. Landis 2017; Landis et al. 2021). It is also important to reiterate that

this model identifies *relationships* but does not determine causality and certainly not direct causality; there may be other underlying causes for the relationships between features and processes, which could be teased apart using a hidden states framework (Beaulieu and O’Meara 2016; Caetano et al. 2018).

One interesting situation we considered was when two included features were highly correlated. This can sometimes lead to model misbehavior, masking the effects of one or both features. In our empirical analysis of *Liolaemus*, we examined the joint RJMCMC probabilities of noneffect for each pair of highly correlated features (categorical and quantitative representations of the same information). This analysis is important; it allows us to incorporate multiple highly correlated sets of parameters without losing inference capabilities. Through marginal probabilities alone, we were able to identify two feature parameters with small probabilities of noneffect: $\sigma_d^{\text{Distance}}$ and ϕ_d^{Distance} . We were also able to examine the other features to determine whether marginal probabilities were being confounded by the duplication of information and identified one additional feature with a low joint probability of noneffect: area with respect to extinction. It is worth comparing the two instances where an effect was inferred.

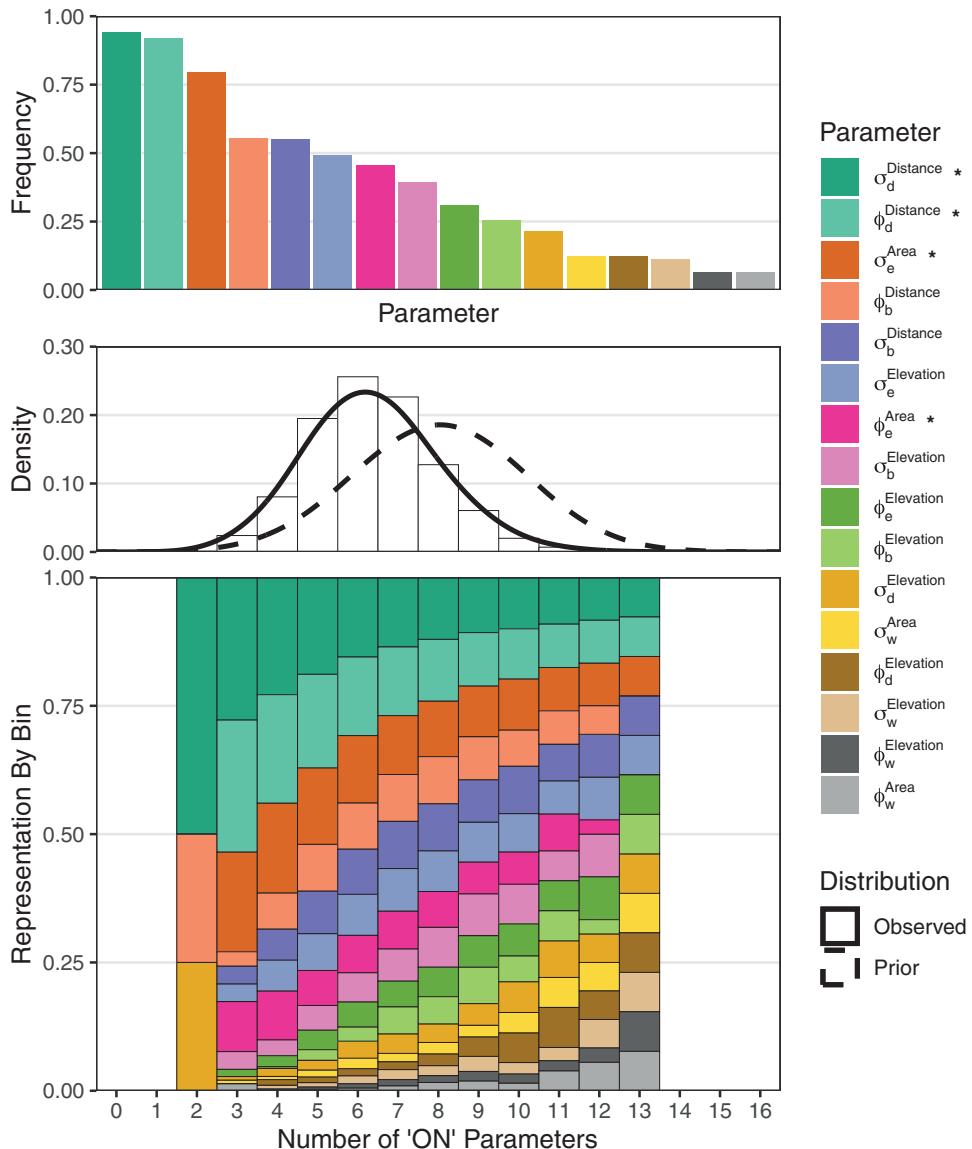


FIGURE 7. Summary of MultiFIG RJMCMC analysis over all considered feature-effect hypotheses. The top panel shows the frequency of being turned “on” for each ϕ and σ parameter in the RJMCMC analysis. The other two panels show a more detailed view; the center panel shows the prior versus posterior distributions for the number of ϕ and σ parameters turned “on” by RJMCMC, and the bottom panel shows the proportional representation of each parameter by bin. Parameters with inferred significance (from individual posteriors and joint reversible-jump probabilities) are marked with an asterisk.

In the first instance (distance), RJMCMC strongly preferred the incorporation of both parameters; in the other instance (size), RJMCMC preferred the incorporation of one parameter, but not necessarily both. While we cannot make any strong statement about the method’s consistency when it comes to selecting between models that represent a feature in a categorical or quantitative way (or both/neither), it is clear that the method does not always prefer more complex models (with multiple parameters representing similar regional feature) or

simpler models (with only one parameter representing similar regional features).

Our model highlights two important feature/process relationships for *Liolaemus*: the relationship between distance and dispersal, and the relationship between area and extinction. A positive value for $\sigma_d^{\text{Distance}}$ and a negative value for ϕ_d^{Distance} indicate that adjacent, closer-together regions have higher rates of dispersal. A negative value for σ_e^{Area} and a negative value for ϕ_e^{Area}

indicate that smaller regions have higher rates of extinction. These relationships are sensible, as distance (Landis et al. 2022) and area-size (Rabosky and Glor 2010; Thomson et al. 2021; Quintero et al. 2023) have previously been implicated in diversification dynamics in other reptile systems using SSE models.

Interestingly, one of the feature/process pairings where our model did not infer significance for either σ or ϕ was elevation and within-region speciation. It was our initial hypothesis that one (or both) of the elevation parameters would show an effect on speciation, since Andean regions possessed higher rates of speciation in the analyses of Esquerre et al. (2019), which examined the family Liolaemidae (including the *Liolaemus* genus). Abiotic (e.g., climatic) and biotic (e.g., competitive) factors are expected to differently shape the ecological niches of otherwise similar lowland versus highland species (Lomolino 2001), which should create different ecological opportunities for speciation and risks for extinction. The outsized historical effects cyclical climate change would have on mountains should further exaggerate differences in net diversification (Flantua et al. 2019; Mastretta-Yanes et al. 2015). That said, our categorization of highland/lowland regions did not initially align with the Andean/non-Andean classification of Esquerre et al.. In our analysis, Patagonia is classified as a "lowland" region, since its mean elevation is below the mean for all regions, but Esquerre et al. classify Patagonia as an Andean region, as the southern part of the Andes falls within Patagonia. Patagonia was thought to be an important source of speciation in the Esquerre et al. analysis. Therefore, we performed a second analysis, replacing our highland/lowland classification with an Andean/non-Andean classification for all six regions following Esquerre et al. (2019). Even with this classification system, our model did not infer significance for parameters associating Andean/non-Andean status with rates of speciation. When we compared relative rates of within-region speciation among regions, Andean regions (in both highland/lowland and Andean/non-Andean analyses) had higher mean within-region speciation rates, even if the model did not indicate that the parameters associated with elevation were significantly different from zero. However, the posterior densities greatly overlapped, and the difference in means was much smaller than Esquerre et al..

We also performed an analysis with a reduced set of features that only assigned regions an Andean or non-Andean classification but did not incorporate areas, distances, or quantitative elevations. This analysis is more closely related to the two-state GeoSSE model that Esquerre et al. used to infer higher rates of speciation in Andean regions but still utilizes six regions. It similarly provided weak support for higher rates of speciation in Andean regions, and parameters associated with this relationship did not show significance, much like in the full analyses with all 16 parameters. In addition, we conducted a 16-parameter analysis using

the highland/lowland classification system with mean elevations replaced by SDs to assess whether elevation variation may be more strongly correlated with within-region speciation; the results of this analysis also indicated no effect (results not shown).

The posterior for the preferred number of parameters offers additional insight into the complex relationships between region features and biogeographic rates. Models with only four parameters were not ultimately preferred, even when no additional parameters could be individually selected. One explanation is that there may be rate variation in the system that is unrelated to the parameters we chose. Additionally, rate variation could be created by some complex interactions of our features which we did not examine, or the model may lack the ability to discern between some correlated parameters (we only examined certain two-parameter joint probabilities). It is also possible that a different choice of priors would lead the model to identify additional significant parameters.

These results highlight the importance of incorporating multiple regional features, both categorical and quantitative, into inferences of biogeographic rates. In addition to recovering differences in speciation rates between Andean and non-Andean regions, we were able to identify important distance/dispersal and area/extinction relationships, which could not have been investigated in a two-region GeoSSE framework. The MultiFIG model can be extended to include several features that are expected to have an affect on different processes and help determine which of those features actually possess a strong link to particular processes.

Our simulation analyses and empirical example made use of a relatively simple model—we incorporated many features, but each feature was assumed to act independently upon each process. This highlights the usefulness of MultiFIG for identifying features of potential importance from a large set of possible features. However, MultiFIG is also useful for more detailed hypothesis testing. To clarify individual feature-process relationships for *Liolaemus*, we performed several analyses with reduced feature sets. Within the MultiFIG framework, it is also possible to construct priors that structure the relationships between feature effects or to customize the functions which relate features to processes (equations 1–8). Because MultiFIG is implemented in RevBayes, analyses are conducted using scripts that define relationships between model components, making the model customizable for those who are looking to answer more specific questions. However, we strongly suggest researchers begin with less restrictive models. If a model returns a counterintuitive result, this is important information; there may be a real biological explanation, or the model/data may not be sufficient for answering that question.

It is also important to note that due to data availability, our analysis of *Liolaemus* used a single fixed,

dated phylogeny. While there seemed to be no important effect of uncertainty on rate estimates under a time-constant FIG model with *Anolis* lizards (Landis et al. 2022), this may not be generalizable to every clade, especially a young clade such as *Lioleamus* with substantial evidence of introgression and incomplete lineage sorting (Esquerre et al. 2021). For researchers who are interested in applying MultiFIG to clades where posterior distributions of trees are available, we currently recommend performing identical analyses on a sample of trees and comparing results. However, developers of RevBayes (Höhna et al. 2019) and TensorPhylo (May and Meyer 2022) are working to enable the joint estimation of divergence times and tree topology with models such as MultiFIG. This will not only incorporate phylogenetic uncertainty into inferences of biogeographic rates and feature effects but will also allow biogeography to inform phylogenetic relationships and divergence times.

This raises another important challenge. As many regional features change over time, we should expect that historical rates of biogeographic processes might also change in kind. However, the FIG model currently makes use of only present-day regional feature data. When analyzing clades of extant taxa, previous work from our collaborations has found that ignoring paleogeography often has little effect on simpler biogeographic parameter estimation or model selection tasks (Landis 2017; Landis et al. 2021). We suspect that this is because the highest proportion statistical information for powering models comes from living species and the regions they currently inhabit, with little coming from historical inferences phylogenetically extrapolated from those extant taxa. But we believe that the reassurances end there because, not surprisingly, we have repeatedly found that ignoring or mischaracterizing paleogeography misleads divergence time and ancestral state estimates buried deeper in the tree (Landis 2017; Landis et al. 2018, 2021). To address this gap, future work will expand the FIG framework to allow for regional features to vary over time, enabling it to model a wider range of hypotheses concerning historical relationships between regional features and biogeographic rates, evolving biome affinities among lineages (Crisp et al. 2009; Lichter-Marck and Baldwin 2023), fossil taxon occurrence and preservation (Xing et al. 2016; Saupe et al. 2019; Landis 2021), and phylogenetic divergence times (Landis 2021).

CONCLUSION

Here we describe the MultiFIG model, which provides a conceptual and mathematical framework for investigating the relationships between multiple regional features and the biogeographic processes of within- and between-region speciation, extinction, and dispersal. Through reversible-jump Markov chain Monte Carlo

and Bayesian parameter estimation, MultiFIG allows for the exploration of models with any number of feature effects, and generates an estimate for the strength and direction of each feature-process relationship. We validate this model through simulation, demonstrating its performance in scenarios with varying numbers of regions, different tree sizes, and several common types of model misspecification. We also apply the MultiFIG model to the South American lizard genus *Lioleamus*, examining 16 regional features related to size, distance, and elevation. We uncover a negative relationship between distance and dispersal, and a negative relationship between area and extinction. These findings attest to the value of incorporating multiple regional features into biogeographic hypothesis testing.

ACKNOWLEDGMENTS

We are grateful to members of the Landis and Myers labs at Washington University in Saint Louis and the Tello lab at the Missouri Botanical Garden for feedback. This work was also enriched by ongoing collaborations with Drs. Isaac Lichter-Marck and Felipe Zapata at UCLA.

SUPPLEMENTARY MATERIAL

Data available from the Dryad Digital Repository: <https://doi.org/10.5061/dryad.dz08kps2x>

FUNDING

This study was supported by the National Science Foundation (NSF) through the Graduate Research Fellowship Program awarded to S.K.S. (NSF DGE 2139839) and through a research grant awarded to M.J.L. (NSF DEB 2040347). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

DATA AVAILABILITY

Data and/or supplementary material and/or software code associated with this paper are available for review via Dryad. The following is a temporary direct download link: <https://doi.org/10.5061/dryad.dz08kps2x>

REFERENCES

- Adriaensen F., Chardon J.P., De Blust G., Swinnen E., Villalba S., Gulnick H., Matthysen E. 2003. The application of 'least-cost' modelling as a functional landscape model. *Landsc. Urban Plan.* 64(4): 233–247.
- Albert J.S., Schoolmaster D.R. Jr., Tagliacollo V., Duke-Sylvester S.M. 2017. Barrier displacement on a neutral landscape: toward a theory of continental biogeography. *Syst. Biol.* 66(2):167–182.

- Alencar L.R.V., Schwery O., Gade M.R., Domínguez-Guerrero S.F., Tarimo E., Bodensteiner B.L., Uyeda J.C., Muñoz M.M. 2024. Opportunity begets opportunity to drive macroevolutionary dynamics of a diverse lizard radiation. *Evol. Lett.* 8(5):623–637.
- Beaulieu J.M., O'Meara B.C. 2016. Detecting hidden diversification shifts in models of Trait-Dependent speciation and extinction. *Syst. Biol.* 65(4):583–601.
- Boschman L.M. 2021. Andean mountain building since the late Cretaceous: a paleoelevation reconstruction. *Earth-Sci. Rev.* 220: 103640.
- Buckley L.B., Davies T.J., Ackerly D.D., Kraft N.J.B., Harrison S.P., Anacker B.L., Cornell H.V., Damschen E.I., Grytnes J.-A., Hawkins B.A., McCain C.M., Stephens P.R., Wiens J.J. 2010. Phylogeny, niche conservatism and the latitudinal diversity gradient in mammals. *Proc. Biol. Sci.* 277(1691):2131–2138.
- Caetano D.S., O'Meara B.C., Beaulieu J.M. 2018. Hidden state models improve state-dependent diversification approaches, including biogeographical models. *Evolution* 72(11):2308–2324.
- Cardillo M., Weston P.H., Reynolds Z.K.M., Olde P.M., Mast A.R., Lemmon E.M., Lemmon A.R., Bromham L. 2017. The phylogeny and biogeography of hakea (proteaceae) reveals the role of biome shifts in a continental plant radiation. *Evolution* 71(8):1928–1943.
- Connor E.F., McCoy E.D. 1979. The statistics and biology of the species-area relationship. *Am. Nat.* 113(6):791–833.
- Crisp M.D., Arroyo M.T.K., Cook L.G., Gandolfo M.A., Jordan G.J., McGlone M.S., Weston P.H., Westoby M., Wilf P., Linder P.H. 2009. Phylogenetic biome conservatism on a global scale. *Nature* 458 (7239):754–756.
- Crisp M.D., Trewick S.A., Cook L.G. 2011. Hypothesis testing in biogeography. *Trends Ecol. Evol.* 26(2):66–72.
- Domínguez-Guerrero S.F., Esquerre D., Burress E.D., Maciel-Mata C.A., Alencar L.R.V., Muñoz M.M. 2024. Viviparity imparts a macroevolutionary signature of ecological opportunity in the body size of female *Liolaemus* lizards. *Nat. Commun.* 15(1): 4966.
- Esquerre D., Brennan I.G., Catullo R.A., Torres-Pérez F., Keogh J.S. 2019. How mountains shape biodiversity: the role of the andes in biogeography, diversification, and reproductive biology in South America's most species-rich lizard radiation (Squamata: Liolaemidae). *Evolution* 73(2):214–230.
- Esquerre D., Keogh J.S., Demangel D., Morando M., Avila L.J., Sites J.W., Ferri-Yáñez F., Leaché A.D. 2022. Rapid radiation and rampant reticulation: Phylogenomics of South American *Liolaemus* lizards. *Syst. Biol.* 71(2):286–300.
- Faria N., Suchard M., Rambaut A., Streicker D., Lemey P. 2013. Simultaneously reconstructing viral cross-species transmission history and identifying the underlying constraints. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 368:20120196. doi: 10.1098/rstb.2012.0196.
- Fine P.V.A., Ree R.H. 2006. Evidence for a time-integrated species-area effect on the latitudinal gradient in tree diversity. *Am. Nat.* 168(6): 796–804.
- Flantua S.G.A., O'Dea A., Onstein R.E., Giraldo C., Hooghiemstra H. 2019. The flickering connectivity system of the north Andean páramos. *J. Biogeogr.* 46(8):1808–1825.
- Freyman W.A., Höhna S. 2017. Cladogenetic and anagenetic models of chromosome number evolution: a Bayesian model averaging approach. *Syst. Biol.* 67(2):195–215.
- Garzione C.N., Hoke G.D., Libarkin J.C., Withers S., MacFadden B., Eiler J., Ghosh P., Mulch A. 2008. Rise of the Andes. *Science* 320 (5881):1304–1307.
- Goldberg E.E., Lancaster L.T., Ree R.H. 2011. Phylogenetic inference of reciprocal effects between geographic range evolution and diversification. *Syst. Biol.* 60(4):451–465.
- Green P.J. 1995. Reversible jump markov chain monte carlo computation and bayesian model determination. *Biometrika* 82(4): 711.
- Green P.J. 2003. Trans-dimensional Markov chain Monte Carlo. In: Green P.J., Hjort N.L., Richardson S., editors. *Highly structured stochastic systems*. Oxford, UK: Oxford University Press. p. 179–198.
- Gregory-Wodzicki K.M. 2000. Uplift history of the Central and Northern Andes: a review. *Geol. Soc. Am. Bull.* 112(7):1091–1105.
- Höhna S., Heath T.A., Boussau B., Landis M.J., Ronquist F., Huelsenbeck J.P. 2014. Probabilistic graphical model representation in phylogenetics. *Syst. Biol.* 63:753–771.
- Höhna S., Landis M.J., Heath T.A., Boussau B., Lartillot N., Moore B.R., Huelsenbeck J.P., Ronquist F. 2016. RevBayes: Bayesian phylogenetic inference using graphical models and interactive model-specification language. *Syst. Biol.* 65:726–736.
- Höhna S., Freyman W.A., Nolen Z., Huelsenbeck J.P., May M.R., Moore B.R. 2019. A Bayesian approach for estimating branch-specific speciation and extinction rates. In Review.
- Huelsenbeck J.P., Larget B., Alfaro M.E. 2004. Bayesian phylogenetic model selection using reversible jump Markov chain Monte Carlo. *Mol. Biol. Evol.* 21:1123–1133.
- Ibargüengoytía N.R., Medina M., Laspur A., Qu Y.-F., Augusto Ramirez Peralta C., Sinervo B., Miles D.B. 2021. Looking at the past to infer into the future: thermal traits track environmental change in *Liolaemidae*. *Evolution* 75(10):2348–2370.
- Landis M.J. 2017. Biogeographic dating of speciation times using paleogeographically informed processes. *Syst. Biol.* 66(23): 128–144.
- Landis M.J. 2021. Biogeographic dating of phylogenetic divergence times using priors and processes. In: Ho S.Y.W., editor. *The molecular evolutionary clock: theory and practice*. Cham, Switzerland: Springer Nature Switzerland. p. 135–155.
- Landis M., Edwards E.J., Donoghue M.J. 2021. Modeling phylogenetic biome shifts on a planet with a past. *Syst. Biol.* 70(1):86–107.
- Landis M.J., Freyman W.A., Baldwin B.G. 2018. Retracing the Hawaiian silversword radiation despite phylogenetic, biogeographic, and paleogeographic uncertainty. *Evolution* 72(11):2343–2359.
- Landis M.J., Matzke N.J., Moore B.R., Huelsenbeck J.P. 2013. Bayesian analysis of biogeography when the number of areas is large. *Syst. Biol.* 62:789–804.
- Landis M.J., Quintero I., Muñoz M.M., Zapata F., Donoghue M.J. 2022. Phylogenetic inference of where species spread or split across barriers. *Proc. Natl. Acad. Sci. U. S. A.* 119(13):e2116948119.
- Lemey P., Rambaut A., Bedford T., Faria N.R., Bielejec F., Baele G., Russell C.A., Smith D.J., Pybus O.G., Brockmann D., Suchard M.A. 2012. The seasonal flight of influenza: a unified framework for spatiotemporal hypothesis testing. *arXiv*.
- Lichter-Marck I.H., Baldwin B.G. 2023. Edaphic specialization onto bare, rocky outcrops as a factor in the evolution of desert angiosperms. *Proc. Natl. Acad. Sci.* 120(6):e2214729120.
- Lomolino M.V. 2001. Elevation gradients of species-density: historical and prospective views. *Glob. Ecol. Biogeogr.* 10(1):3–13.
- MacArthur R.H., Wilson E.O. 1967. *The theory of island biogeography*. revised ed. NJ, USA: Princeton University Press.
- Maddison W.P., Midford P.E., Otto S.P. 2007. Estimating a binary character's effect on speciation and extinction. *Syst. Biol.* 56(5): 701–710.
- Mastretta-Yanes A., Moreno-Letelier A., Piñero D., Jorgensen T.H., Emerson B.C. 2015. Biodiversity in the mexican highlands and the interaction of geology, geography and climate within the transmexican volcanic belt. *J. Biogeogr.* 42(9):1586–1600.
- Matzke N.J. 2014. Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. *Syst. Biol.* 63:951–970.
- May M.R., Meyer X. 2022. Tensorphylo.
- NASA. 2013. Shuttle radar topography mission (SRTM) global.
- Navas C.A. 2002. Herpetological diversity along andean elevational gradients: links with physiological ecology and evolutionary physiology. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* 133(3):469–485.
- Quintero I., Landis M.J. 2020. Interdependent phenotypic and biogeographic evolution driven by biotic interactions. *Syst. Biol.* 69(4): 739–755.
- Quintero I., Landis M.J., Jetz W., Morlon H. 2023. The build-up of the present-day tropical diversity of tetrapods. *Proc. Natl. Acad. Sci.* 120 (20):e2220672120.
- Rabosky D.L., Glor R.E. 2010. Equilibrium speciation dynamics in a model adaptive radiation of island lizards. *Proc. Natl. Acad. Sci.* 107 (51):22178–22183.

- Ree R.H., Sanmartín I. 2009. Prospects and challenges for parametric models in historical biogeographical inference. *J. Biogeogr.* 36(7): 1211–1220.
- Ree R.H., Sanmartín I. 2018. Conceptual and statistical problems with the dec+ j model of founder-event speciation and its comparison with dec via model selection. *J. Biogeogr.* 45(4):741–749.
- Ree R.H., Smith S.A. 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Syst. Biol.* 57(1):4–14.
- Ree R.H., Moore B.R., Webb C.O., Donoghue M.J. 2005. A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution* 59(11):2299–2311.
- Ronquist F. 1997. Dispersal-Vicariance analysis: a new approach to the quantification of historical biogeography. *Syst. Biol.* 46(1): 195–203.
- Sanmartín I., Van Der Mark P., Ronquist F. 2008. Inferring dispersal: a Bayesian approach to phylogeny-based island biogeography, with special reference to the canary islands. *J. Biogeogr.* 35(3): 428–449.
- Saupe E.E., Myers C.E., Townsend Peterson A., Soberón J., Sin-garayer J., Valdes P., Qiao H. 2019. Spatio-temporal climate change contributes to latitudinal diversity gradients. *Nat. Ecol. Evol.* 3(10): 1419–1429.
- Scheiner S.M., Rey-Benayas J.M. 1994. Global patterns of plant diversity. *Evol. Ecol.* 8:331–347.
- Skeels A., Cardillo M. 2018. Reconstructing the geography of speciation from contemporary biodiversity data. *Am. Nat.* 193(2):240–255.
- Skeels A., Esquerre D., Lipsky D., Pellissier L., Boschman L.M. 2023. Elevational goldilocks zone underlies the exceptional diversity of a large lizard radiation (*Liolaemus*; *Liolaemidae*). *Evolution*. 77(12):2672–2686.
- Thomson R.C., Spinks P.Q., Shaffer H.B. 2021. A global phylogeny of turtles reveals a burst of climate-associated diversification on continental margins. *Proc. Natl. Acad. Sci.* 118(7):e2012215118.
- Valente L.M., Phillimore A.B., Etienne R.S. 2015. Equilibrium and non-equilibrium dynamics simultaneously operate in the Galápagos islands. *Ecol. Lett.* 18(8):844–852.
- Xing Y., Gandolfo M.A., Onstein R.E., Cantrill D.J., Jacobs B.F., Jordan G.J., Lee D.E., Popova S., Srivastava R., Su T., Vikulin S.V., Yabe A., Linder H.P. 2016. Testing the biases in the rich cenozoic angiosperm macrofossil record. *Int. J. Plant Sci.* 177(4):371–388.